
Coupled Semiempirical Quantum Mechanics and Molecular Mechanics (QM/MM) Calculations on the Aqueous Solvation Free Energies of Ionized Molecules

PETER L. CUMMINS, JILL E. GREADY

Division of Biochemistry and Molecular Biology, John Curtin School of Medical Research, Australian National University, P.O. BOX 334, Canberra, ACT 2601, Australia

Received 13 May 1998; accepted 24 February 1999

ABSTRACT: The aqueous solvation free energies of ionized molecules were computed using a coupled quantum mechanical and molecular mechanical (QM/MM) model based on the AM1, MNDO, and PM3 semiempirical molecular orbital methods for the solute molecule and the TIP3P molecular mechanics model for liquid water. The present work is an extension of our model for neutral solutes where we assumed that the total free energy is the sum of components derived from the electrostatic/polarization terms in the Hamiltonian plus an empirical "nonpolar" term. The electrostatic/polarization contributions to the solvation free energies were computed using molecular dynamics (MD) simulation and thermodynamic integration techniques, while the nonpolar contributions were taken from the literature. The contribution to the electrostatic/polarization component of the free energy due to nonbonded interactions outside the cutoff radii used in the MD simulations was approximated by a Born solvation term. The experimental free energies were reproduced satisfactorily using variational parameters from the vdW terms as in the original model, in addition to a parameter from the one-electron integral terms. The new one-electron parameter was required to account for the short-range effects of overlapping atomic charge densities. The radial distribution

Correspondence to: J. E. Gready; e-mail: jill.gready@anu.edu.au

Contract/grant sponsor: Australian National University (Strategic Development Grant)

functions obtained from the MD simulations showed the expected H-bonded structures between the ionized solute molecule and solvent molecules. We also obtained satisfactory results by neglecting both the empirical nonpolar term and the electronic polarization of the solute, i.e., by implementing a nonpolarization model. © 1999 John Wiley & Sons, Inc. J Comput Chem 20: 1028–1038, 1999

Keywords: quantum mechanical/molecular mechanical; solvation; free energy; electrostatics; polarization

Introduction

Semiempirical quantum mechanics (QM) and molecular mechanics (MM) methods are characterized by the use of empirically derived parameters. In each of these methods, the canonical equations contain parameters that are varied to give a best fit to a range of known properties. To produce a useful model that couples quantum mechanics and molecular mechanics (QM/MM), additional parameterization is required. There are now several QM/MM approaches based on the semiempirical AM1, MNDO, and PM3 molecular orbital methods reported in the literature.^{1–4} However, there remains considerable scope for improving the quantitative accuracy of the coupled semiempirical QM and MM method. In previous work⁴ we proposed a QM/MM model where parameter values in the equation for the free energy of solvation obtained from molecular dynamics (MD) simulation protocols were chosen to reproduce the experimental free energy values. Although requiring much larger amounts of computation time, this MD simulation approach provided a more accurate model for condensed-phase calculations compared with the other QM/MM models that rely on data for small clusters or gas-phase bimolecular complexes.^{1,2} The other significant difference between the two types of approaches was in the treatment of electrostatic interactions between QM and MM atoms. The QM/MM one-electron terms in our model required no new parameters, but were given by the coulomb potential in the semiempirical approximation.^{5,6} Only particular parameters within the vdW terms for interactions between the quantum and molecular mechanics atoms were modified.

The parameter sets for our original model were derived for neutral molecules in aqueous solution,

whereas Vasilyev et al.² used complexes formed between ionized and neutral species in the gas phase. However, we did not report the results of ion solvation because we found that both our original model and the model of Vasilyev et al.² were generally incapable of accurately predicting free energies. Adding a correction for the truncation of the electrostatic interactions in the MD simulations is essential where ions are involved, but this does not solve the problem, as in many cases the solvation free energies are already far too negative compared with experiment. In such instances where the solvation free energies are overestimated compared with the experiment, increasing the vdW radii for interactions between the QM solute molecule and the MM solvent molecules reduces the free energy but at the expense of the H-bonding between the ion and solvent. This suggests the problem lies in the description of the electrostatic interactions over short distances between the ion and solvent molecules.

In this article we develop our QM/MM model further to include ionized species. The model was parameterized using a set of molecules⁷ that are analogues of the charged side chains of the ionizable amino acids. To correct the deficiencies in the description of the short-range electrostatic interactions we adopted the one-electron terms suggested by Vasilyev et al.² Although there are several ways to treat long-range electrostatic interactions in MD simulations, for reasons of computational simplicity, we used the Born approximation⁸ for the free energy contribution due to electrostatic interactions outside the cutoff radii. In the usual implementation of the QM/MM method, the electronic charge distribution of the QM part of the system is polarized by the electrostatic potential of the MM part. However, it is certainly possible to perform the QM/MM calculations neglecting the QM polarization. In the present study we also considered this nonpolarization option for the calculation of the solvation free energy.

Methods

QM/MM MODEL

The total potential energy of the system partitioned into quantum (solute) and molecular mechanics (solvent) groups of atoms is given by

$$E_T = E_{QM} + E_{MM} + E_{QM/MM} \quad (1)$$

where E_{QM} is the energy of the quantum system, E_{MM} is the energy of the molecular mechanics part of the system, and $E_{QM/MM}$ is the interaction energy between the quantum and molecular mechanics parts of the system. In the present model, E_{QM} is computed using the semiempirical AM1,^{9,10} MNDO,^{11,12} and PM3¹³ methods. The QM/MM interaction energy is the sum of polar (electrostatic/polarization) and nonpolar (van der Waals) terms,

$$E_{QM/MM} = (1 - \lambda)E_{ele} + E_{vdW} \quad (2)$$

where we have introduced the parameter λ for scaling the electrostatic/polarization term in the MD simulations to determine the solvation free energy. Note that in the present model, polarization refers only to the electronic charge density of the QM solute: no explicit electronic polarization is included for the MM solvent.

The two-center one-electron integrals $I_{\nu\mu}$ in AM1, MNDO, and PM3 are given by two-center two-electron repulsion terms of the form (ss, $\nu\mu$).¹⁴ If R is the distance between atoms A and B,

$$I_{ss} = (R^2 + \alpha_0^2)^{-1/2} \quad (3a)$$

$$I_{sz} = \frac{1}{2} \left[\left\{ (R - D_1)^2 + \alpha_1^2 \right\}^{-1/2} - \left\{ (R + D_1)^2 + \alpha_1^2 \right\}^{-1/2} \right] \quad (3b)$$

$$I_{zz} = I_{ss} + \frac{1}{4} \left[\left\{ (R + 2D_2)^2 + \alpha_2^2 \right\}^{-1/2} + \left\{ (R - 2D_2)^2 + \alpha_2^2 \right\}^{-1/2} - \frac{1}{2} (R^2 + \alpha_2^2)^{-1/2} \right] \quad (3c)$$

$$I_{pp} = I_{ss} + \frac{1}{2} \left[(R^2 + 4D_2^2 + \alpha_2^2)^{-1/2} - (R^2 + \alpha_2^2)^{-1/2} \right] \quad (3d)$$

where

$$\alpha_l^2 = (\rho_l^A + \rho_l^B)^2 \quad (4)$$

In eqs. (3) and (4) ρ_l and D_l are atomic parameters. Note that the core-core repulsions are given by integrals of type I_{ss} . These types of one-electron terms are also required for the calculation of E_{ele} . However, in the QM/MM method of Vasilyev et al.² eq. (4) becomes

$$\alpha_l^2 = \sigma (\rho_l^{QM} + \rho_l^{MM})^2 \quad (5)$$

where σ is an empirical scaling parameter with values of 0.095 for the AM1 and MNDO methods and 0.097 for the PM3 method. These values were the result of a fit to data on the gas-phase ion-molecule complexes. In our model calculations on polar and nonpolar solute molecules in solution, these terms were set equal to zero, i.e., $\sigma = 0$, and eq. (3) simplifies to⁶

$$I_{ss} = 1/R \quad (6a)$$

$$I_{sz} = \frac{1}{2} \left[(R - D_1)^{-1} - (R + D_1)^{-1} \right] \quad (6b)$$

$$I_{zz} = \frac{1}{2} \left[\frac{1}{2} \left\{ (R + 2D_2)^{-1} + (R - 2D_2)^{-1} \right\} + I_{ss} \right] \quad (6c)$$

$$I_{pp} = \frac{1}{2} \left[(R^2 + 4D_2^2)^{-1/2} + I_{ss} \right] \quad (6d)$$

Equation (6) is purely coulombic, arising from a distribution of point charges centered on the atoms, while the former eq. (3) gives rise to a screened coulomb potential. The screened potential is required in semiempirical QM methods because the actual charge distribution is continuous. In molecular calculations the charge distributions overlap, leading to a deviation from coulomb behavior. Values of atomic parameters ρ_l are chosen so that in the limit of maximum overlap ($R = 0$) eq. (3) yields the correct one-center integral.¹⁴ In the opposite limit ($R \rightarrow \infty$) both represent pure coulomb potentials: we previously demonstrated that a multipole expansion of the coulombic potential indeed yields the exact dipole moment of a molecule.⁶ Clearly, the importance of ρ_l in determining accurate representations of the one-electron integrals depends on the extent of overlap of atomic charge densities. Consequently, eq. (6) may not adequately describe ionic solutes due to the overlap with the solvent charge densities. This overlap is simply the result of the strong electro-

static interaction between the ion and polar solute molecules. In the case of weaker interactions involving neutral species, the error due to the neglect of overlap is relatively small, and can be readily absorbed into the parameterization of the vdW terms. These errors are larger for ions and may not be corrected using realistic values of the vdW parameters.

POLARIZATION MODEL

The present scheme differs from that used to treat neutral molecules by the introduction of nonzero scaling parameters σ [eqs. (3) and (5)], in addition to the vdW radii parameters, and a Born solvation term to correct for the neglect of solute–solvent electrostatic interactions beyond the cutoff radius used in the MD simulations. Thus, the total solvation free energy (ΔG_{sol}) for ionized molecules can be written as

$$\Delta G_{\text{sol}} = N^{-1} \sum_{i=1}^N [E_{\text{ele}}]_i + \frac{1-\varepsilon}{2\varepsilon r_{\text{cut}}} Q^2 + \Delta G_{\text{c/vdW}} \quad (7)$$

The first term in eq. (7) is readily derived from the equations for thermodynamic integration.^{15,16} It is calculated in a molecular dynamics (MD) simulation by the creation or annihilation of the solute–solvent interaction terms via the λ coupling parameter in eq. (2). This corresponds to either a charging (solvation) or discharging (desolvation) of the solute in water. N is the number of integration time steps, and $[E_{\text{ele}}]_i$ is the electrostatic/polarization part of the QM/MM interaction [eq. (3)] evaluated using the configuration obtained at the i th time step in the MD simulation. The second term is the Born correction, in which ε is the dielectric constant of the solvent, Q is the total charge on the molecule, and r_{cut} is the cutoff radii for the neglect of solute–solvent interactions in the calculation of E_{ele} . The remaining nonpolar (i.e., cavity and van der Waals) free energy term $\Delta G_{\text{c/vdW}}$ is entirely empirical, assuming a linear relationship between free energy and solvent-accessible surface area of the solute as used by Sitkoff et al.⁷

NONPOLARIZATION MODEL

Given the wave function for a system, the total interaction energy between molecular fragments may be variously decomposed into electrostatic,

exchange, polarization, and charge transfer contributions.^{17,18} In QM/MM models the wave function is defined for only the quantum part of the system. The charge distribution of the QM part can be polarized by the electrostatic potential produced by the atom-centered charges q_i of the MM system by performing the self-consistent field (SCF) calculation using the appropriately perturbed one-electron Fock matrix elements.⁴ Clearly, electrostatic and polarization contributions are contained in the first two terms of eq. (7). The effects of exchange and charge transfer are implicit in the remaining $\Delta G_{\text{c/vdW}}$ term. However, electrostatic models without explicit polarization terms are capable of predicting the binding energies, structures, and force constants of molecular complexes, provided that the short-range repulsions are also accurately described.^{19–24} The importance of short-range repulsions is reflected in our choice of vdW parameters in the QM/MM solvation model.⁴ The combined effect of exchange, polarization, and charge transfer on the interaction energy is generally found to be of much lesser importance. Thus, the electrostatic energy by itself can often provide a good estimate of the binding energy.

These considerations lead us to consider for our QM/MM system a simplified nonpolarization model in which only purely electrostatic terms are retained. The solvation free energy for such a model system is then given by

$$\Delta G_{\text{sol}} = N^{-1} \sum_{i=1}^N [E_{\text{ele}}^0]_i + \frac{1-\varepsilon}{2\varepsilon r_{\text{cut}}} Q^2 \quad (8)$$

where E_{ele}^0 is the electrostatic contribution to the solute–solvent interaction energy in the absence of any polarization of the solute's charge density, i.e., a first-order perturbation correction due to the solute–solvent electrostatic interaction. Clearly, we can expect that a nonpolarization model may provide a reasonable description of only the energetic and structural features of a system. However, this is all that is required for computing an MD trajectory and, thus, calculation of the thermodynamic properties of the system. Molecular properties that are sensitive to changes in the molecules charge density cannot, of course, be computed using this model.

LINEAR RESPONSE ASSUMPTION

The coupling approach used to obtain free energies from molecular simulation can be thought of as a charging of the molecule in solution. If the

system (solute plus solvent) response to this charging process is linear, the ratio of the free energy to the average solute-solvent electrostatic/polarization energy is exactly one-half,²⁵⁻²⁷ i.e.,

$$\Delta G_{\text{ele}}/\langle E_{\text{ele}} \rangle = 1/2 \quad (9)$$

where in our model ΔG_{ele} is the free energy corresponding to the first term in eqs. (7) and (8), i.e., the free energy obtained by thermodynamic integration, and E_{ele} is the corresponding electrostatic/polarization energy in eq. (2) for the QM/MM interaction. This linear response assumption has been shown to be approximately true for the aqueous solvation free energy of ions obtained by simulations using MM force fields,²⁶⁻²⁸ suggesting that a reliable estimate of the free energy of a solvation or binding process may be obtained without the need to perform thermodynamic integration or perturbation calculations. This approach for obtaining the free energy may be especially useful in protein-ligand binding studies.²⁷⁻²⁹ We have also computed $\langle E_{\text{ele}} \rangle$ to test the validity of the linear response assumption for the QM/MM approach.

MOLECULAR DYNAMICS SIMULATIONS

The systems consisting of an ion, treated by QM, in a box containing 800 to 1000 TIP3P³⁰ water molecules were initially energy minimized for 100 cycles of conjugate gradients, followed by the MD simulations starting with $\lambda = 0$ in eq. (2). After an initial equilibration phase of MD, the free energies were obtained by integration between $\lambda = 0$ and $\lambda = 1$ over a 60-ps simulation time with λ coupled continuously and linearly to the MD time step.¹⁵

The free energy change was then recalculated by performing the reverse $\lambda = 1$ to $\lambda = 0$ integration over 60 ps. Longer integration times were used for the ammonium ion as a test of convergence. We used the constant temperature and pressure algorithm of Berendsen et al.³¹ to perform the MD simulations, with a time step of 0.001 ps and hydrogen masses set to 3 amu. The temperature was set at 300 K, with a relaxation time of 0.1 ps, and the pressure set at 1 atm. A 9 Å cutoff was used for the neglect of nonbonded solvent-solvent interactions and 12 Å for the solute-solvent interactions. The covalent bonds were unconstrained in the MD simulations. All calculations were performed using Molecular Orbital Programs for Simulations (MOPS).³²

Results and Discussion

POLARIZATION MODEL

The electrostatic shielding parameter σ , and the radii R_e values obtained for the H-bond 10-12 potential,⁴ are given in Table I. Details of the approach used to optimize the parameters are described elsewhere.⁴ Except for the *N*-propylguanidinium ion, the R_e values are unchanged from those determined for neutral molecules in solution.⁴ The vdW radii R^* for the 6-12 potential³³ are also the same, except that the amine nitrogen (N) atoms in the PM3 model increases to 1.85 from 1.70 Å. The nonpolar cavity/vdW contributions ($\Delta G_{c/vdW}$) obtained from Sitkoff et al.,⁷ the AM1 electrostatic/polarization contribution to the free energy ΔG_{ele} obtained by thermodynamic integration, and the corresponding average AM1 en-

TABLE I. Electrostatic Shielding Parameters σ and vdW Parameters R_e (Å) Obtained for Different Types of Ionized Groups in the AM1, MNDO, and PM3 Models with Polarization.

Ionized Group	H-bonding ^a	AM1		MNDO		PM3	
		σ	R_e	σ	R_e	σ	R_e
Ammonium	-H ⁺ ... OW	0.07	1.90	0.08	1.80	0.00	1.60 ^b
Thiol ion	-S ⁻ ... HW	0.06	3.95	0.12	3.40	0.06	3.95
Carboxylate	-O ⁻ ... HW	0.12	2.65	0.10	2.65	0.14	2.65
Phenolate	-O ⁻ ... HW	0.18	1.90	0.23	2.00	0.25	2.10
Imidazolium	-H ⁺ ... OW	0.00	2.15	0.02	2.15	0.00	2.10
Guanidinium	-H ⁺ ... OW	0.00	2.15	0.00	2.00	0.00	1.80

^aOW = water oxygen. HW = water hydrogen.

^b $R^* = 1.85$ Å for vdW radius of the N atom (see text).

TABLE II.

Cavity / vdW ($\Delta G_{c/vdW}$) and AM1 Polarization Model Electrostatic (ΔG_{ele}) Components of the Solvation Free Energies (kcal / mol) of Molecular Ions and the Average Solute – Solvent AM1 Electrostatic Interaction Energies $\langle E_{ele} \rangle$.

Ion	$\Delta G_{c/vdW}^a$	ΔG_{ele}^b	$\langle E_{ele} \rangle$	$\Delta G_{ele} / \langle E_{ele} \rangle$
Ammonium	1.71	-68.29 ± 0.74	-154.79 ± 0.46	0.441
<i>N</i> -butylammonium	2.36	-59.86 ± 0.49	-136.97 ± 0.87	0.437
Methylthiol ion	1.88	-65.38 ± 0.20	-145.52 ± 0.92	0.449
Acetate	1.94	-67.81 ± 0.43	-157.82 ± 0.62	0.430
Propionate	2.11	-67.34 ± 0.67	-157.53 ± 0.93	0.427
<i>p</i> -cresol ion	2.40	-63.10 ± 0.20	-150.51 ± 0.54	0.419
Methylimidazolium	2.21	-52.32 ± 0.25	-112.51 ± 0.45	0.465
<i>N</i> -propylguanidinium	2.46	-54.61 ± 0.11	-120.64 ± 1.54	0.453

^aCavity / vdW contributions from ref. 7.

^bMean and standard deviation of the forward ($\lambda = 0 \rightarrow 1$) and reverse ($\lambda = 1 \rightarrow 0$) integrations.

ergy $\langle E_{ele} \rangle$ and ratios $\Delta G_{ele} / \langle E_{ele} \rangle$ [eq. (9)] are given in Table II. Note that all free energy values are quoted as the mean and standard deviation of the forward ($\lambda = 0 \rightarrow 1$) and reverse ($\lambda = 1 \rightarrow 0$) integrations. The total AM1, MNDO, and PM3 free energies of solvation (ΔG_{sol}) obtained using the polarization model [eq. (7)] are given in Table III, together with the free energies obtained from experiment.^{7,38,39} Using the parameters in Table I, the differences between the AM1, MNDO, and PM3 simulation free energies of solvation and the

experimental results are typically not much greater than 1 kcal/mol, which is similar to the estimated error in the thermodynamic integration. However, larger errors of 3 kcal/mol were obtained for the ammonium and butylammonium ions with the PM3 model. The ratios $\Delta G_{ele} / \langle E_{ele} \rangle$ show that the systems exhibit small but significant deviations from linear behavior.

Note that the training set, i.e., the ions used in the fitting procedure to obtain the QM/MM force-field parameters (Table I), consists of the represen-

TABLE III.

AM1, MNDO, and PM3 Polarization Model Solvation Free Energies ΔG_{sol} (kcal / mol) of Molecular Ions Compared with Experiment.

Ion	AM1 ^a	MNDO ^a	PM3 ^a	Exp.
Ammonium	-80.46 ± 0.74	-80.71 ± 0.01	-84.53 ± 0.43	$-81.53^b, -79^c$
Methylammonium ^f	-73.6 ± 0.5	-71.2 ± 1.1	-74.5 ± 0.7	-70^c
Dimethylammonium ^f	-65.7 ± 1.5	-66.3 ± 0.8	-65.2 ± 1.1	-63^c
Trimethylammonium ^f	-58.2 ± 0.8	-59.0 ± 1.2	-59.0 ± 1.0	-59^c
<i>N</i> -butylammonium	-71.33 ± 0.49	-70.70 ± 0.19	-72.87 ± 0.11	-69.24^b
Methylthiol ion	-77.33 ± 0.20	-77.30 ± 0.38	-76.04 ± 0.06	-76.79^b
Acetate	-79.70 ± 0.43	-81.58 ± 0.50	-78.94 ± 0.55	$-80.65^b, -77^c$
Propionate	-79.06 ± 0.67	-80.55 ± 0.41	-78.63 ± 0.37	-79.12^b
Benzoate ^f	-81.9 ± 1.6	-83.9 ± 1.6	-82.8 ± 1.5	-76^d
<i>p</i> -cresol ion	-74.53 ± 0.20	-75.52 ± 0.59	-74.96 ± 0.34	-75.01^b
Phenol ion ^f	-76.8 ± 1.1	-76.5 ± 1.2	-76.9 ± 0.9	-72^c
Methanol ion ^f	-92.2 ± 0.8	-79.1 ± 0.8	-87.4 ± 1.1	-95^c
Methylimidazolium	-63.94 ± 0.25	-63.81 ± 0.32	-65.76 ± 0.32	-64.13^b
Pyridinium ^f	-59.6 ± 0.8	-61.3 ± 0.7	-59.9 ± 0.9	-59^c
<i>N</i> -propylguanidinium	-65.98 ± 0.11	-66.88 ± 0.67	-64.30 ± 0.48	-66.07^e

^aMean and standard deviation of the forward ($\lambda = 0 \rightarrow 1$) and reverse ($\lambda = 1 \rightarrow 0$) integrations.

^bExperimental values as quoted in ref. 7 used for training set of ions (see text).

^cExperimental values from ref. 38.

^dEstimate, ref. 39.

^ePoisson–Boltzman calculation from ref. 7.

^fTest set of ions (see text), assuming $\Delta G_{c/vdW} = 2.0 \pm 0.4$ kcal / mol (Table II).

tatives of the amino acid side chains. This set is rather small compared with the one used in our studies of neutral solutes.⁴ Consequently, in order to validate the underlying physical model, we performed free energy calculations on a "test" set of ions not included in this training set. Overall, the test set results (Table III) are quite satisfactory. In particular, the trend in the reduction of the free energy on successive methylation of the ammonium ion is well reproduced. Also, when parameters derived for $\text{N}-\text{H}^+$ in a five-member heterocycle (imidazolium) are transferred to a six-member ring (pyridinium) the agreement with experiment remains very good. The free energy calculated for the phenol ion using parameters derived for the *p*-cresol ion is also in good agreement with the experiment. However, if we transfer these parameters to the methanol ion, we find that, while correctly predicting larger solvation free energies than for the phenol and *p*-cresol ions, the values are somewhat underestimated compared with the experiment. Similarly, if we attach an aromatic ring to the carboxylate group (benzoate), rather than an alkyl chain (acetate and propionate), we find a consistently small but significant overestimation of the free energy compared with an experimental estimate.³⁹ These errors do not point to a deficiency in the model, but simply reflect the fact that the chemical environment (i.e., aromatic or alkyl) of the functional group is sufficiently different to require a revised set of parameter values.⁴ Clearly, many other ionized functional group types (e.g., $-\text{OH}_2^+$, $-\text{SH}_2^+$) would also require a unique set of parameter values. In the current work the types have been restricted to those of amino acid side chains required for our purpose of simulations of proteins. However, the advantage of our functional-group QM/MM approach is that the model is readily extendable to any type of chemical species for which there exists an example of an experimental free energy.

As a further test of the error size in the calculation of ΔG_{ele} we also carried out integrations over longer time scales for the ammonium ion. Figure 1 shows the variations in the mean of the total free energies (ΔG_{sol}) obtained from the forward and reverse integrations over simulation times ranging from 5 ps to 1 ns in both directions. The error bars indicate the standard deviation from the mean in each of the calculations. In the limit where the integration proceeds infinitely slowly, the work needed to drive the change is reversible, and the free energy is the same in both directions. Nevertheless, over finite times useful upper and lower

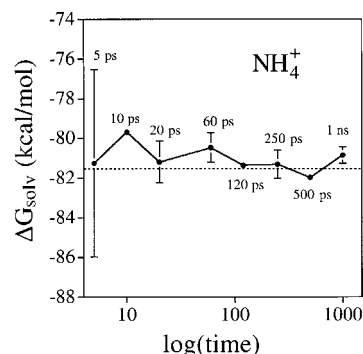


FIGURE 1. AM1 model solvation free energy [polarization model, eq. (7)] of the ammonium ion calculated using 5, 10, 20, 60, 120, 250, 500, and 1000 ps of simulation time in both the forward and reverse thermodynamic integrations. The mean free energies obtained from the forward and reverse integrations are plotted as a \log_{10} function of the simulation times in ps. The hysteresis for each of the simulation times is indicated by the error bars. The horizontal dashed line indicates the experimental value for the solvation free energy.

bounds for the free energy may be obtained.³⁴ For simulation times as short as 5 ps we obtained a mean free energy estimate within 1 kcal/mol of experiment. These differences represent no more than a few percent of the total solvation free energy.

To validate the Born correction, the solvation free energies of the ammonium and acetate ions were calculated using different choices for the solute-solvent and solvent-solvent cutoff radii. The results are given in Table IV. The effect of simulation box size, i.e., the number of water molecules, on the computed free energies was negligible. However, a larger box size is required to eliminate boundary artefacts when larger cutoff radii are used. The differences in the calculated free energy on increasing the solute-solvent radius from 12 to 18 Å are -3 kcal/mol for the ammonium ion and -2 kcal/mol for the acetate ion. The effect of the solvent-solvent cutoff on solvation free energies of ions has been noted and discussed in detail elsewhere in the literature.^{16,35-37} Briefly, the significant effect is that the interaction energy between the ion and solvent water will always be too large when solvent-solvent cutoff radii are used. Consequently, as the cutoff radius between water molecules is increased the artefact is removed and the solute-solvent interaction energy becomes lower. In the present model, increasing the solvent-solvent radius from 9 to 12 Å reduces the

TABLE IV.

 AM1 Polarization Model Free Energies ΔG_{sol} (kcal/mol) of Selected Ions with Various Cutoff Radii R_{cut} (Å) for the Neglect of Nonbonded Interactions.

Ion	R_{cut}		Solvent Number	ΔG_{sol}^b
	(solute-solvent) ^a	(solvent-solvent)		
Ammonium	12	9	815	-80.46 ± 0.74
	12	9	2321	-81.88 ± 1.43
	18	9	2321	-84.96 ± 0.51
	18	12	2321	-77.07 ± 0.07
Acetate	12	9	921	-79.70 ± 0.43
	12	9	2520	-79.99 ± 0.70
	18	9	2520	-81.81 ± 0.18
	18	12	2520	-75.77 ± 0.52

^a $R_{\text{cut}} = 12$ Å; Born correction = -13.83 kcal/mol. $R_{\text{cut}} = 18$ Å; Born correction = -9.22 kcal/mol.

^bMean and standard deviation of the forward ($\lambda = 0 \rightarrow 1$) and reverse ($\lambda = 1 \rightarrow 0$) integrations.

solvation free energy by almost 8 kcal/mol for the ammonium ion and 6.5 kcal/mol for the acetate ion. Although these effects are quite large, they are partly compensated for by the error in the Born solvation term so that the magnitude of the computed solvation free energies fall 4 to 5 kcal/mol below experiment. Thus, where simulations are performed using larger cutoff radii, values of σ that are somewhat smaller than those reported in Table I may be more appropriate.

Radial distribution functions, $g(r)$, for the solute-solvent H-bond interactions are given in Figures 2 to 4. The expected H-bond distances are reproduced quite well. The $\text{O}^- \cdots \text{HW}$ and $\text{H}^+ \cdots \text{OW}$ distances range between 1.6 to 2.0 Å, while the $\text{O}^- \cdots \text{OW}$ and $\text{N} \cdots \text{OW}$ distances are in the 2.5 to 2.9 Å range. The $\text{S}^- \cdots \text{HW}$ distances are 2.8 Å for AM1 and PM3 and 2.4 Å for MNDO. These results agree well with the QM/MM calculations of Vasilyev et al.² on gas-phase ion-water complexes. For most of the possible H-bonding sites more than one H-bond is predicted. The exceptions are for the $\text{H}^+ \cdots \text{OW}$ bond in the *N*-propylguanidinium ion where there are ca. 50% (PM3) and 80% (MNDO) probabilities of H-bonding. The AM1 and MNDO $\text{H}^+ \cdots \text{OW}$ interactions between the ammonium ion and solvent exhibits a very prominent second peak at 3.2 Å. However, PM3 seems to produce a broadening of these two distinct peaks into a single plateau region between 2 and 4 Å. Similarly, the PM3 H-bond peak is rather poorly resolved for the *N*-propylguanidinium ion.

NONPOLARIZATION MODEL

The AM1 electrostatic free energy components ΔG_{ele} obtained from integration and the corre-

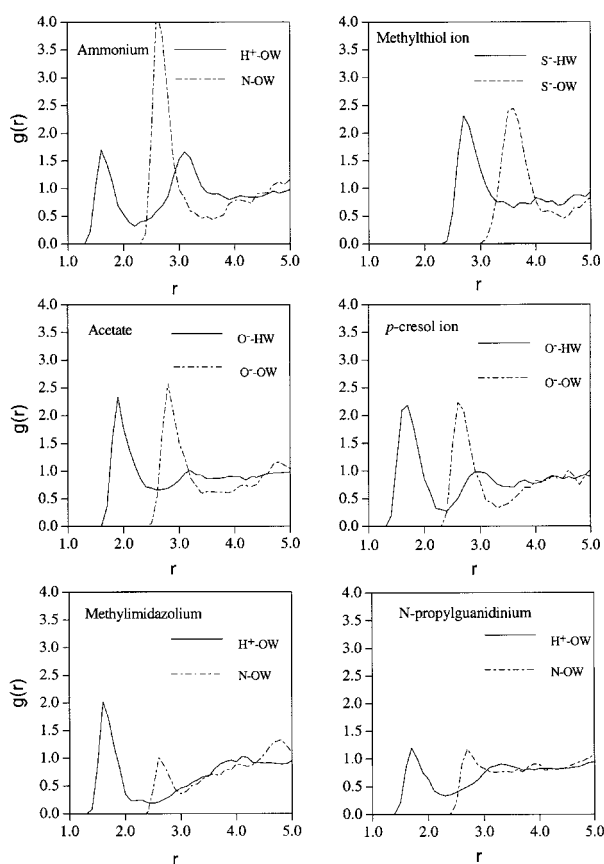


FIGURE 2. AM1 polarization model (Table III) radial distribution functions $g(r)$ for the solute-solvent H-bond interactions.

sponding average AM1 electrostatic energies $\langle E_{\text{ele}} \rangle$ for the nonpolarization model are given in Table V. The solute polarization terms ΔG_{pol} , as given by the difference between free energies obtained from the polarization and nonpolarization models, are

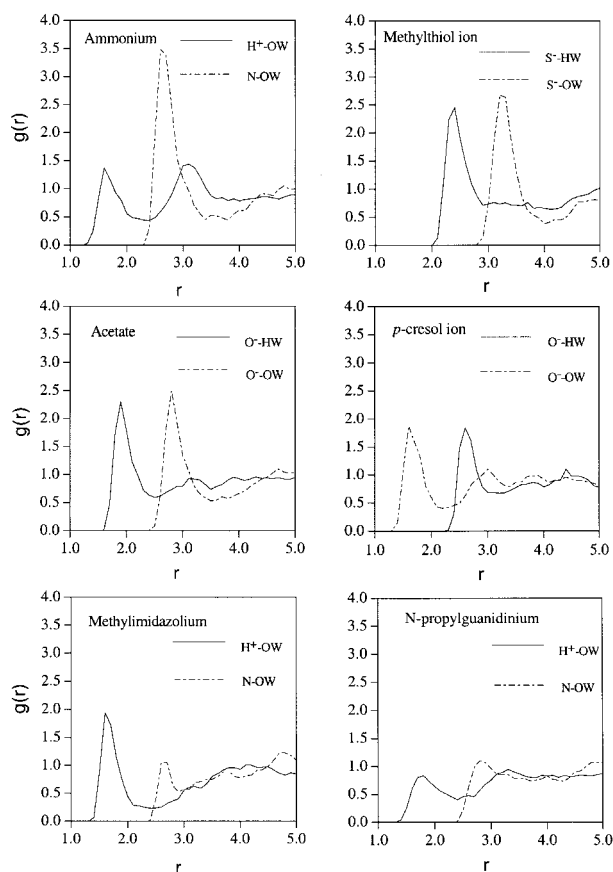


FIGURE 3. MNDO polarization model (Table III) radial distribution functions $g(r/\text{\AA})$ for the solute-solvent H-bond interactions.

also given in Table V. The total free energies of solvation (ΔG_{sol}) for the AM1, MNDO, and PM3 methods obtained using the nonpolarization model [eq. (8)] are given in Table VI, together with the solvation free energies obtained from experiment.⁷ The parameter values used were those derived for the polarization model calculations (Table I). The polarization terms, ΔG_{pol} , are relatively small, and tend to cancel the nonelectrostatic corrections ($\Delta G_{\text{c/vdW}}$) listed in Table II. Consequently, the nonpolarization model free energies of solvation are in reasonably good agreement with the experimental values. The polarization terms range from almost negligible for the ammonium and *N*-propylguanidinium ions up to a maximum of -4.2 kcal/mol for the *p*-cresol ion. With the exception of the ammonium, butylammonium, and *p*-cresol ions, where the errors are ca. 3 kcal/mol, the differences between the calculated and experimental free energies are within 2 kcal/mol. Although small, the solute polarization term is highly non-

linear. The ratios $\Delta G_{\text{ele}}/\langle E_{\text{ele}} \rangle$ for the nonpolarization model range in value from 0.441 to 0.496 compared with the lower valued range of 0.419 to 0.465 for the polarization model.

Conclusions

We have modified our coupled semiempirical (AM1, MNDO, and PM3) and molecular mechanics (TIP3P) model to include ionized organic molecules in aqueous solution. The parameterization space of the model includes both the electronic and vdW terms for the interaction between the QM and MM parts of the system, in particular, those terms that are likely to have a major influence on H-bonding. The force-field parameters required different values, depending on the type of functional group involved in H-bonding and the QM method (AM1, MNDO, or PM3) used to model the solute molecule. The model also provided a reasonable description of solute-solvent H-bonded

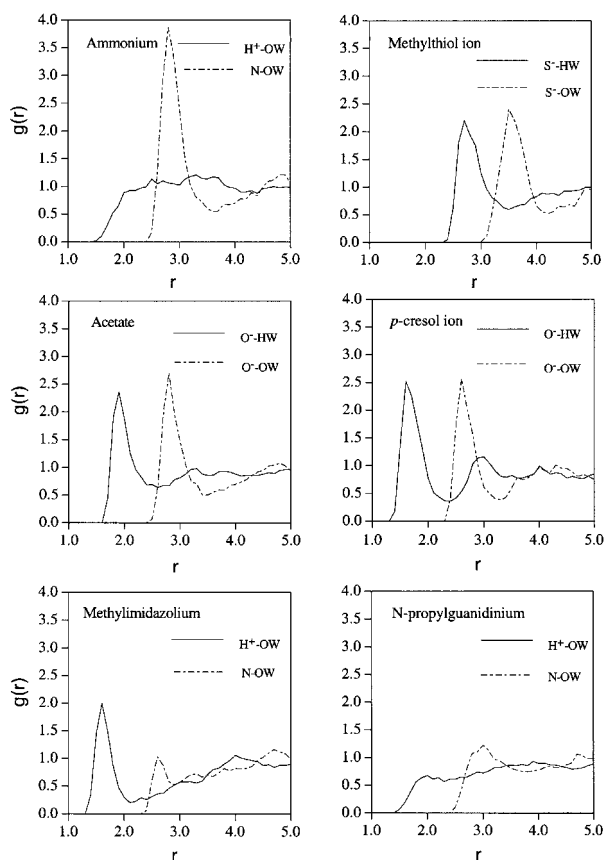


FIGURE 4. PM3 polarization model (Table III) radial distribution functions $g(r/\text{\AA})$ for the solute-solvent H-bond interactions.

TABLE V. **AM1 Polarization (ΔG_{pol}) and Nonpolarization Model Electrostatic (ΔG_{ele}) Components of the Solvation Free Energies (kcal / mol) of Molecular Ions and the Average AM1 Nonpolarization Model Solute – Solvent Electrostatic Interaction Energies $\langle E_{\text{ele}} \rangle$.**

Ion	$\Delta G_{\text{pol}}^{\text{a}}$	$\delta G_{\text{ele}}^{\text{b}}$	$\langle E_{\text{ele}} \rangle$	$\Delta G_{\text{ele}} / \langle E_{\text{ele}} \rangle$
Ammonium	-0.20 ± 1.19	-68.09 ± 0.45	-154.41 ± 0.71	0.441
<i>N</i> -butylammonium	-3.94 ± 0.78	-55.92 ± 0.29	-123.95 ± 0.48	0.451
Methylthiol ion	-1.07 ± 0.90	-64.31 ± 0.70	-139.76 ± 0.28	0.460
Acetate	-2.78 ± 1.02	-65.03 ± 0.59	-145.06 ± 0.71	0.448
Propionate	-3.60 ± 1.12	-63.74 ± 0.45	-143.45 ± 1.04	0.444
<i>p</i> -cresol ion	-4.16 ± 0.22	-58.94 ± 0.02	-128.96 ± 0.63	0.457
Methylimidazolium	-2.44 ± 1.20	-49.88 ± 0.95	-104.43 ± 0.35	0.478
<i>N</i> -propylguanidinium	$+0.06 \pm 0.51$	-54.66 ± 0.40	-110.28 ± 0.88	0.496

^a $\Delta G_{\text{pol}} = \Delta G_{\text{ele}}(\text{polarization model}) - \Delta G_{\text{ele}}(\text{nonpolarization model})$.

^bMean and standard deviation of the forward ($\lambda = 0 \rightarrow 1$) and reverse ($\lambda = 1 \rightarrow 0$) integrations.

structures. The neglect of electronic polarization of the QM part of the system by the MM atomic charges had a small effect on the free energy. The linear response approximation holds better for the nonpolarization model. Thus, the mean interaction energies may give more reliable estimates of free energy changes [eq. (9)] within the nonpolarization model.

In general, we found that the treatment of ions requires a screened coulomb potential to account for the effects of overlapping atomic charge densities. In our model, originally developed for neutral solutes,⁴ we did not explicitly consider the effect of charge overlap, and adopted a purely coulombic potential for the interaction. It may be possible to develop one set of screening (σ) parameters applicable to both ionized and neutral species. How-

ever, this would certainly require at least a partial reparameterization of the vdW terms, considerably more computational effort than in the approach we have taken, and may not necessarily yield better overall accuracy. We plan to use the present QM/MM model to study protein–ligand interactions, particularly in the rationalization of enzymic reaction mechanisms.³⁸

Acknowledgments

We gratefully acknowledge the Australian National University Supercomputer Facility for generous grants of computer time, and the support of an ANU Strategic Development Grant.

TABLE VI. **AM1, MNDO, and PM3 Nonpolarization Model Solvation Free Energies ΔG_{sol} (kcal / mol) [eq. (8)] of Molecular Ions Compared with Experiment.**

Ionized Molecule	AM1 ^a	MNDO ^a	PM3 ^a	Exp. ^b
Ammonium	-81.92 ± 0.45	-81.96 ± 0.55	-85.58 ± 0.39	-81.53
<i>N</i> -butylammonium	-69.75 ± 0.29	-70.74 ± 0.70	-73.30 ± 0.96	-69.24
Methylthiol ion	-78.14 ± 0.70	-77.39 ± 0.26	-75.96 ± 0.83	-76.79
Acetate	-78.86 ± 0.59	-78.09 ± 0.10	-80.11 ± 0.50	-80.44
Propionate	-77.57 ± 0.45	-77.16 ± 0.06	-79.19 ± 0.70	-79.12
<i>p</i> -cresol ion	-72.77 ± 0.02	-71.29 ± 0.72	-72.17 ± 0.41	-75.01
Methylimidazolium	-63.71 ± 0.95	-64.44 ± 0.53	-63.88 ± 0.63	-64.13
<i>N</i> -propylguanidinium	-68.49 ± 0.40	-65.93 ± 0.23	-64.17 ± 0.54	-66.07^{c}

^aMean and standard deviation of the forward ($\lambda = 0 \rightarrow 1$) and reverse ($\lambda = 1 \rightarrow 0$) integrations.

^bExperimental values as quoted in ref. 7.

^cPoisson–Boltzman calculation from ref. 7.

References

1. Field, M. J.; Bash, P. A.; Karplus, M. *J Comput Chem* 1990, 11, 700.
2. Vasilyev, V. V.; Blyznyuk, A. A.; Voityuk, A. A. *Int J Quantum Chem* 1992, 44, 897.
3. Bash, P. A.; Ho, L. L.; MacKerell, A. D.; Levine, D.; Hallstrom, P. *Proc Natl Acad Sci USA* 1996, 93, 3698.
4. Cummins, P. L.; Gready, J. E. *J Comput Chem* 1997, 18, 1496.
5. Chudinov, G. E.; Napolov, D. V.; Basilevsky, M. U. *Chem Phys* 1992, 160, 41.
6. Cummins, P. L.; Gready, J. E. *Chem Phys Lett* 1994, 225, 11.
7. Sitkoff, D.; Sharp, K. A.; Honig, B. *J Phys Chem* 1994, 98, 1978.
8. Born, M. *Z Phys* 1920, 1, 45.
9. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J Am Chem Soc* 1985, 107, 3902.
10. Dewar, M. J. S.; Yuan, Y.-C. *Inorg Chem* 1990, 129, 3881.
11. Dewar, M. J. S.; Thiel, W. *J Am Chem Soc* 1977, 99, 4899.
12. Dewar, M. J. S.; Reynolds, C. H. *J Comput Chem* 1986, 7, 140.
13. Stewart, J. J. P. *J Comput Chem* 1989, 10, 209.
14. Dewar, M. J. S.; Thiel, W. *Theor Chim Acta* 1977, 46, 89.
15. Mezei, M.; Beveridge, D. L. *Ann NY Acad Sci* 1986, 1, 482.
16. Cummins, P. L.; Gready, J. E. *J Comput Chem* 1994, 15, 704.
17. Morokuma, K. *J Chem Phys* 1971, 55, 1236.
18. Kitaura, K.; Morokuma, K. *Int J Quantum Chem* 1976, 10, 325.
19. Buckingham, A. D.; Fowler, P. W. *J Chem Phys* 1983, 79, 6426.
20. Buckingham, A. D.; Fowler, P. W. *Can J Chem* 1985, 63, 6426.
21. Rendell, A. P. L.; Bacskey, G. B.; Hush, N. S. *Chem Phys Lett* 1985, 117, 400.
22. Hurst, G. J. B.; Fowler, P. W.; Stone, A. J.; Buckingham, A. D. *Int J Quantum Chem* 1986, 29, 1223.
23. Cummins, P. L.; Rendell, A. P. L.; Swanton, D. J.; Bacskey, G. B.; Hush, N. S. *Int Rev Phys Chem* 1986, 5, 139.
24. Spackman, M. A. *J Chem Phys* 1986, 85, 6587.
25. Warshel, A.; Russell, S. T. *Q Rev Biophys* 1984, 17, 283.
26. Roux, B.; Yu, H.-A.; Karplus, M. *J Phys Chem* 1990, 94, 4683.
27. Åqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng* 1994, 7, 385.
28. Åqvist, J. *J Comput Chem* 1996, 17, 1587.
29. Gorse, A.-D.; Gready, J. E. *Protein Eng* 1997, 10, 23.
30. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J Chem Phys* 1983, 79, 926.
31. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J Chem Phys* 1984, 81, 3684.
32. Cummins, P. L. *Molecular Orbital Programs for Simulations (MOPS)*, 1996.
33. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J Comput Chem* 1986, 7, 230.
34. Hunter, J. E., III; Reinhardt, W. P.; Davis, T. *J Chem Phys* 1993, 99, 6856.
35. Straatsma, T. P.; Berendsen, H. J. C. *J Chem Phys* 1988, 89, 5876.
36. Åqvist, J. *J Phys Chem* 1990, 94, 8021.
37. Åqvist, J. *J Phys Chem* 1994, 98, 8253.
38. Pearson, R. G. *J Am Chem Soc* 1986, 108, 6109.
39. Cramer, C. J.; Truhlar, D. G. *Rev Comput Chem* 1995, 6, 1.
40. Cummins, P. L.; Gready, J. E. *J Comput Chem* 1998, 19, 977.